PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07H 21/02, C07K 1/00, G01N 33/53, C12P 21/06, C12N 5/00, A61K 39/395, A01K 67/00

A1

(11) International Publication Number: **WO 00/52026**

(43) International Publication Date: 8 September 2000 (08.09.00)

(21) International Application Number: PCT/US00/05712

(22) International Filing Date: 3 March 2000 (03.03.00)

(30) Priority Data:

60/122,612	3 March 1999 (03.03.99)	US
60/153,235	13 September 1999 (13.09.99)	US

(71) Applicant *(for all designated States except US)*: GENE LOGIC, INC. [US/US]; 708 Quince Orchard Road, Gaithersburg, MD 20878 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SUN, Hong-Wei [US/US]; 10419 Dalebrook Lane, Potomac, MD 20854 (US). MUNGER, William [US/US]; 6 Ardmour Court, Bethesda, MD 20816 (US). AKIYAMA, Kiyotaka [JP/JP]; 6-1, Zengyo 1-chome, Fujisawa, Kanagawa (JP).

(74) Agent: TUSCAN, Michael, S.; Morgan, Lewis & Bockius LLP, 1800 M Street, NW, Washington, DC 20036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: IDENTIFICATION OF A cDNA ASSOCIATED WITH COMPENSATORY HYPERTROPHY IN RENAL TISSUE

(57) Abstract

The invention relates generally to the changes in gene expression in kidney tissue undergoing rapid growth in a renal compensatory hypertrophy model. The invention relates specifically to a novel human gene which corresponds to a mRNA which is up-regulated in the compensatory hypertrophy model.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDENTIFICATION OF A cDNA ASSOCIATED WITH COMPENSATORY HYPERTROPHY IN RENAL TISSUE

INVENTORS: Hong-Wei Sun, Kiyotaka Akiyama and William E. Munger

FIELD OF THE INVENTION

The invention relates generally to the changes in gene expression in kidney tissue induced during renal compensatory hypertrophy. The invention relates specifically to a novel human gene which is expressed in kidney tissue. This application is claims priority to Provisional Application 60/122,612, filed March 3, 1999 and Provisional Application 60/153,235, filed September 13, 1999, both of which are herein incorporated by reference in their entirety.

BACKGROUND OF THE INVENTION

Abnormal renal function is among the most common ailments requiring intensive medical care. In addition, the incidence and prevalence rates of end-stage renal disease (ESRD) in the United States continue to increase. In 1995, the incidence rate was 262 per million population, with a point prevalence rate of 975 per million population. The exact number of individuals with abnormal renal function but who have not yet progressed to ESRD is difficult to assess. The incidence rate per million population of treated end-stage renal disease (ESRD) has been increasing at similar rates in most countries that record counts of new ESRD patients per year. Data from the United States Renal Data System (USRDS) suggest an exponential growth for both incidence rates and prevalence rates (Port FK, *Dis Mon* 1998 May;44(5):214-234).

Diabetic nephropathy is the most common cause of end-stage renal disease (ESRD) and accounts for 35% of the ESRD population in the United States. It results in considerable morbidity, mortality, and expense. The average cost of managing one diabetic patient with ESRD is approximately $50,000 a year (Kobrin SM, *Kidney Int Suppl* 1997 Dec;63:S144-150). Approximately 5.8 million people in the United States have been

diagnosed by a physician as being diabetic, and an additional 4 to 5 million people have undiagnosed diabetes. Although the incidence of new cases of diabetes appears to be declining from a peak of 300 per 100,000 population in 1973, to 230 per 100,000 in 1981, its prevalence continues to rise, due to a 19 percent decline since 1970 in deaths caused by diabetes. In 1982, 34,583 deaths were attributed to diabetes, resulting in diabetes being ranked as the seventh leading underlying cause of death. Medical and surgical complications of diabetes due to macro- and microvascular disease result in 5,800 new cases of blindness, 4,500 perinatal deaths, 40,000 lower extremity amputations and 3,000 deaths due to diabetic coma (ketotic and hyperosmolar) and at least 4,000 new cases of end-stage renal disease (Verh K, *Acad Geneeskd Belg* 1989;51(2):81-149; discussion 149-151).

A characteristic and early manifestation of diabetes in humans is renal hypertrophy (Bakris GL *et al., Dis Mon* (1993) 39(8):573-611). This compensation mechanism is a physiological response in which the cells of the kidney increase in size and protein content without synthesizing DNA or dividing (Kujubu *et al., Am J Physiol* (1991) 260(6 Pt 2): F823-827). The compensatory growth of the kidney is a highly regulated process (Wolf *et al., Kidney Int* (1991) 39(3): 401-420). Early studies demonstrated significant increases in ribosome synthesis, including increased rDNA transcription rates (Ouellette *et al., Am J Physiol* (1987) 253(4 Pt 1): C506-513). Subsequent studies have demonstrated the induction of early response genes such as the *TIS* family; proto-oncogenes *ras*, *fos*, *myc*; structural protein genes such as vimentin and β-actin; and transport protein genes (Na^+, K^+ -ATPase, ADP-ATP translocase and calcyclin (Norman *et al., Proc Natl Acad Sci* USA (1988) 85(18): 6768-6772; Moskowitz *et al., J Urol* (1995) 154(4): 1560-1565; Sawczuk *et al., J Urol* (1988) 140(5 Pt 2: 1145-1148; Kujubu *et al., Am J Physiol* (1991) 260(6 Pt 2): F823-827; Beer *et al., J Cell Physiol* (1987) 131(1): 29-35). Renal hypertrophy involves a gradual and progressive increase in mRNA levels resulting in the coordinated expression of positive and negative growth control elements. Such sustained message amplification observed during renal compensation represents a distinct cellular response which is distinguishable from the responses that characterize several pathways of cell differentiation (*e.g.*, proto-oncogene expression patterns in hyperplasia in liver).

If intervention is expected to be successful in halting or slowing down renal disease progression, means of accurately assessing the early manifestations of renal

disease need to be established. One way to accurately assess the early manifestations of renal disease is to identify markers which are uniquely associated with disease progression. Likewise, the development of therapeutics to prevent or repair kidney damage relies on the identification of kidney genes responsible for kidney cell growth and function.

SUMMARY OF THE INVENTION

The present invention is based on the discovery of a new gene which is expressed during renal compensatory hypertrophy. The invention includes isolated nucleic acid molecules selected from the group consisting of an isolated nucleic acid molecule that encodes the human amino acid sequence of SEQ ID NO: 2, an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NO: 2, an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO: 1 and an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2. Nucleic acid molecules of the invention may have 60% nucleotide sequence identity through the open reading frame of SEQ ID NO: 1, preferably about 70-75% sequence identity, more preferably about 80-85% sequence identity, and even more preferably at least about 90% sequence identity through the open reading frame. Nucleic acid molecules of the invention also include the above described molecules derived from or that encode the murine amino acid sequence of SEQ ID NO: 4.

The present invention further includes the nucleic acid molecules operably linked to one or more expression control elements, including vectors comprising the isolated nucleic acid molecules. The invention further includes host cells transformed to contain the nucleic acid molecules of the invention and methods for producing a protein comprising the step of culturing a host cell transformed with a nucleic acid molecule of the invention under conditions in which the protein is expressed.

The invention further provides an isolated polypeptide selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2, an isolated polypeptide comprising a functional fragment of at least 10 amino acids of SEQ ID NO: 2, an isolated polypeptide comprising conservative amino acid substitutions

of SEQ ID NO: 2 and naturally occurring amino acid sequence variants of SEQ ID NO: 2. Polypeptides of the invention also include polypeptides with an amino acid sequence having at least about 50%, 60%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 2 more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% sequence identity with the sequence set forth in SEQ ID NO: 2. Polypeptides as described above derived from the murine amino acid sequence of SEQ ID NO: 4 are also included.

The invention further provides an isolated antibody that specifically binds to a polypeptide of the invention, including monoclonal and polyclonal antibodies.

The invention further provides methods of identifying an agent which modulates the expression of a nucleic acid encoding a protein of the invention, comprising the steps of: exposing cells which express the nucleic acid to the agent; and determining whether the agent modulates expression of said nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

The invention further provides methods of identifying an agent which modulates at least one activity of a protein of the invention, comprising the steps of: exposing cells which express the protein to the agent; and determining whether the agent modulates at least one activity of said protein, thereby identifying an agent which modulates at least one activity of the protein.

The invention further provides methods of identifying binding partners for a protein of the invention, comprising the steps of: exposing said protein to a potential binding partner; and determining if the potential binding partner binds to said protein, thereby identifying binding partners for the protein.

The present invention further provides methods of modulating the expression of a nucleic acid encoding a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein. The invention also provides methods of modulating at least one activity of a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates at least one activity of the protein.

The present invention further includes non-human transgenic animals modified to contain the nucleic acid molecules of the invention or mutated nucleic acid molecules such that expression of the polypeptides of the invention is prevented.

The invention further provides methods of diagnosing renal or other disease states, comprising the steps of: acquiring a tissue or blood or other sample from a subject; and determining the level of expression of nucleic acid molecules of the invention or polypeptides of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 Figure 1 shows the alignment of the amino acid sequences of SEQ ID NO: 2 (upper sequence, human) and SEQ ID NO: 4 (lower sequence, murine).

Figure 2 Figure 2 shows plots predicting the hydrophilic regions of the amino acid sequence of a) SEQ ID NO: 2, and b) SEQ ID NO: 4. More positive values indicate hydrophobicity. Analysis using Kyte-Doolittle is shown as a black line, for comparison with the analysis of Goldman et al shown as a grey line.

Figure 3 Figure 3 shows the tissue distribution of RNA encoding the protein of SEQ ID NO: 2 as analyzed by Northern blot in human 1) heart, 2) brain, 3) placenta, 4) lung, 5) liver, 6) muscle, 7) kidney and 8) pancreas. Lane M contains an RNA marker.

Figure 4 Figure 4 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined in various tissues by quantitative PCR analysis with 20 (left panel) or 25 PCR cycles (right panel). M is a 100 bp DNA Marker; Lane 1) brain; 2) heart; 3) skeletal muscles; 4) kidney; 5) leukocytes; 6) liver; 7) lung; 8) spleen.

Figure 5 Figure 5 shows the relative expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined by quantitative PCR analysis in diseased and control kidney tissues, relative to the expression of GAPDH. "Min. Ch." refers to minimum change; "IgAN" refers to IgA nephropathy; "NCGN" refers to necrotizing crescentic glomerular nephropathy.

Figure 6 Figure 6 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined by quantitative PCR analysis in various tissues relative to the expression of GAPDH.

Figure 7 Figure 7 shows READS-based differential display to measure levels of a nucleic acid sequence of the invention in a model of renal compensatory hypertrophy. In this model, one kidney is surgically removed, resulting in the remaining kidney

undergoing compensatory hypertrophy over subsequent months. Animals were sacrificed at various time points and differential gene expression between the remaining kidney in the nephrectomized animals vs. controls was evaluated by READS. The two time points shown demonstrate that a band on the READS gel corresponding to levels of mRNA encoding a nucleic acid of the invention are upregulated in the remaining kidney from nephrectomized animals undergoing compensatory hypertrophy but not in control animals.

Figure 8 Figure 8 shows A) the growth activity and B) the expression levels of mRNA encoding a protein related to SEQ ID NO: 2, in rat mesangial cells at the indicated number of hours following serum removal (to generate conditions of growth arrest induced by serum starvation). "Log" indicates cells are in log phase (i.e., controls for growth arrest).

Figure 9 Figure 9 shows A) the growth activity and B) the expression levels of mRNA encoding a protein related to SEQ ID NO: 2 in rat mesangial cells at the indicated times following the re-initiation of growth activity. Treatment consisted of growth arrest induced by growth factor (GF) deprivation followed by serum addition to re-initiate growth. "Log" indicates cells are in log phase (i.e., controls for growth arrest).

Figure 10 Figure 10 shows the induction of GN-01 expression by TNF-α treatment of cells. The expression levels of mRNA encoding the protein of SEQ ID NO: 2 (designated "GN-01") in human umbilical vein endothelial cells (HuVec) and human umbilical artery endothelial cells (HuAec) were determined by quantitative PCR relative to the expression of GAPDH. Cells were treated with TNFα at time 0 in a time course study.

Figure 11 Figure 11 shows A) an analysis of cell cycle and cell death effects on rat mesangial cells transfected with a clone that encodes for the mRNA encoding the protein of SEQ ID NO: 2 (designated "GN-01"), or transfected with a similar mutated clone (designated "GN-01*"), or transfected with a LacZ control vector. Transfection with the GN-01 clone resulted in enhanced cell death while the mutant GN-01 clone, GN-01* was less able to induce death. B) Cell cycle and cell death analysis was performed using flow cytometry and measuring BrdU uptake as detected by anti-BrdU-FITC antibodies and by standard gating methods.

Figure 12 Figure 12 shows that transfection of rat mesangial cells with a clone that encodes for the mRNA encoding the protein of SEQ ID NO: 2 results in a functional shift in the cell cycle compared with LacZ transfected controls such that cells spend less

time in S phase and more time in G2, indicating that a cell cycle block occurs in G2. Transfection of the rat mesangial cells under the same conditions with the GN-01* mutant (containing a single amino acid substitution (L->P) within the leucine zipper motif) results in no significant changes in the cell cycle compared with LacZ transfected controls. Cell cycle analysis was performed using flow cytometry to measure DNA content (with Propridium Iodide), expression of GN-01 with anti-FLAG-FITC labeled antibodies, and the application of standard gating techniques.

Figure 13 In Figure 13, confocal microscopy images show the localization of the protein of SEQ ID NO: 2 (designated "GN-01") that has been exogenously expressed in rat mesangial cells (b and d), contrasted with images of the same cell that reveal gross cellular structure (a and c). Both nuclear (b and d) and non-nuclear (d) localization was observed, indicating that subcellular localization is a regulated process, likely related to the functional role of the protein of SEQ ID NO: 2.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

I. General Description

The present invention is based in part on identifying a new gene that is expressed in both human and murine kidney tissue and which corresponds to a murine EST which is up-regulated in a renal compensatory hypertrophy model. The human gene encodes a protein predicted to consist of 291 amino acids (also 291 amino acids in the mouse).

The protein can serve as a target for agents that modulate its expression or activity. For example, agents may be identified which modulate biological processes associated with renal disease, kidney transplantation and kidney regeneration.

The present invention is further based on the development of methods for isolating binding partners that bind to the protein. Probes based on the protein are used as capture probes to isolate potential binding partners, such as other proteins. Dominant negative proteins, DNAs encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. Additionally, the protein provides a novel target for the screening of synthetic small molecules and combinatorial or naturally occurring compound libraries to discover novel

therapeutics to regulate kidney function.

II. Specific Embodiments

A. The Protein Associated with Kidney Growth

The present invention provides isolated protein, allelic variants of the protein, and conservative amino acid substitutions of the protein. As used herein, the "protein" or "polypeptide" refers, in part, to a protein that has the human or mouse amino acid sequence depicted in SEQ ID NO: 2 or SEQ ID NO: 4. The terms also refer to naturally occurring allelic variants and proteins that have a slightly different amino acid sequence than that specifically recited above. Allelic variants, though possessing a slightly different amino acid sequence than those recited above, will still have the same or similar biological functions associated with the 291 amino acid protein.

As used herein, the family of proteins related to the human amino acid sequence of SEQ ID NO: 2 refers to proteins that have been isolated from organisms in addition to humans. The methods used to identify and isolate other members of the family of proteins related to the 291 amino acid protein, such as the mouse protein comprising the sequence of SEQ ID NO: 4, are described below.

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein.

The proteins of the present invention further include insertion, deletion or conservative amino acid substitution variants of the sequences set forth in SEQ ID NO: 2 or SEQ ID NO: 4. As used herein, a conservative variant refers to alterations in the amino acid sequence that do not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least 75% amino acid sequence identity with the sequences set forth in SEQ ID NO: 2 or SEQ ID NO: 4, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95%. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity (see section B for the relevant parameters). Fusion proteins, or N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

Thus, the proteins of the present invention include molecules having the amino acid sequence disclosed in SEQ ID NO: 2 or SEQ ID NO: 4; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues of either the human or murine protein; amino acid sequence variants wherein one or more amino acid residues has been inserted N- or C-terminal to, or within, the disclosed coding sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by another residue. Such fragments, also referred to as peptides or polypeptides, may contain antigenic regions, functional regions of the protein identified as regions of the amino acid sequence which correspond to known protein domains, as well as regions of pronounced hydrophilicity. The regions are all easily identifiable by using commonly available protein sequence analysis software such as MacVector™ (Oxford Molecular). See also Figure 2 and Example 2.

Contemplated variants further include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

As described below, members of the family of proteins can be used: 1) to identify agents which modulate at least one activity of the protein; 2) to identify binding partners for the protein, 3) as an antigen to raise polyclonal or monoclonal antibodies, and 4) as a therapeutic agent or target.

B. Nucleic Acid Molecules

The present invention further provides nucleic acid molecules that encode the protein having SEQ ID NO: 2 or SEQ ID NO: 4 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to such a nucleic acid and remains stably bound to it under appropriate stringency conditions, or encodes a polypeptide sharing at least about 50%, 60%, 70% or 75% sequence identity, preferably at least about 80%, and more preferably at least about 85%, and even more preferably at least about 90 or 95% or more identity with the peptide sequences. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

Homology or identity at the nucleotide or amino acid sequence level is determined by **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) analysis using the algorithm employed by the programs **blastp, blastn, blastx, tblastn** and **tblastx** (Karlin , et al. Proc. Natl. Acad. Sci. USA 87: 2264-2268 (1990) and Altschul, S. F. J. Mol. Evol. 36: 290-300(1993), fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the **BLAST** program is to first consider similar segments between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul et al. (Nature Genetics 6: 119-129 (1994)) which is fully incorporated by reference. The search parameters for **histogram**,

descriptions, **alignments**, **expect** (i.e., the statistical significance threshold for reporting matches against database sequences), **cutoff**, **matrix** and **filter** are at the default settings. The default scoring matrix used by **blastp**, **blastx**, **tblastn**, and **tblastx** is the **BLOSUM62** matrix (Henikoff, et al. Proc. Natl. Acad. Sci. USA 89: 10915-10919 (1992), fully incorporated by reference). For **blastn**, the scoring matrix is set by the ratios of **M** (i.e., the reward score for a pair of matching residues) to **N** (i.e., the penalty score for mismatching residues), wherein the default values for **M** and **N** are 5 and -4, respectively. Four **blastn** parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent **Blastp** parameter settings were Q=9; R=2; wink=1; and gapw=32. A **Bestfit** comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C., or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C Another example is use of 50% formamide, 5 x SSC (0.75M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 1 or SEQ ID NO: 3 and which encode a functional protein. Preferred hybridizing molecules are those that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO: 1 or SEQ ID NO: 3.

As used herein, a nucleic acid molecule is said to be "isolated" when the nucleic acid molecule is substantially separated from contaminant nucleic acid molecules encoding other polypeptides.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. For instance, fragments which encode peptides corresponding to predicted antigenic regions may be prepared (see Figure 2 and Example 2). If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming (see the discussion in Section H).

Fragments of the encoding nucleic acid molecules of the present invention (i.e., synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention, can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, (*J. Am. Chem. Soc.* **103**:3185-3191, 1981) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can readily employ any such label to obtain labeled variants of the nucleic acid molecules of the invention.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

C. Isolation of Other Related Nucleic Acid Molecules

As described above, the identification and characterization of the nucleic acid

molecule having SEQ ID NO: 1 or SEQ ID NO: 3 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the protein family in addition to the sequences herein described. Further, the presently disclosed nucleic acid molecules allow a skilled artisan to isolate nucleic acid molecules that encode other members of the family of proteins in addition to the proteins having SEQ ID NO: 2 or SEQ ID NO: 4.

Essentially, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as lambda gtll library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the enzyme.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

D. rDNA molecules Containing a Nucleic Acid Molecule

The present invention further provides recombinant DNA molecules (rDNAs) that contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in situ*. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook *et al.,Molecular Cloning: A Laboratory Manual.* 2nd Ed. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1989. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control

sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, e.g., protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from BioRad Laboratories, (Richmond, CA), pPL and pKK223 available from Pharmacia (Piscataway, NJ).

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form rDNA molecules that contain a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction

sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), the vector pCDM8 described herein, and the like eukaryotic expression vectors.

Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, i.e., the neomycin phosphotransferase (*neo*) gene. (Southern *et al.*, *J. Mol. Anal. Genet.* **1**:327-341, 1982.) Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

E. Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells (NIH/3T3) available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like eukaryotic tissue culture cell lines.

Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with a rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed, see, for example, Cohen *et al.*, *Proc. Natl. Acad. Sci. USA* **69**:2110, 1972; and Maniatis *et al.*, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982). With

regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for example, Graham *et al.*, *Virol.* **52**:456, 1973; Wigler *et al.*, *Proc. Natl. Acad. Sci. USA* **76**:1373-76, 1979.

Successfully transformed cells, i.e., cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern, *J. Mol. Biol.* **98**:503, 1975, or Berent *et al.*, *Biotech.* **3**:208, 1985 or the proteins produced from the cell assayed via an immunological method.

F. Production of Recombinant Proteins using a rDNA Molecule

The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

First, a nucleic acid molecule is obtained that encodes a protein of the invention, such as the nucleic acid molecule depicted in SEQ ID NO: 1 or SEQ ID NO: 3, nucleotides 257-1129 of SEQ ID NO: 1, nucleotides 257-1132 of SEQ ID NO: 1, nucleotides 130-1002 of SEQ ID NO: 3, nucleotides 130-1005 of SEQ ID NO: 3 or a nucleic acid molecule which encodes SEQ ID NO: 5. If the encoding sequence is uninterrupted by introns, it is directly suitable for expression in any host.

The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is

accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

G. Methods to Identify Binding Partners

Another embodiment of the present invention provides methods for use in isolating and identifying binding partners of proteins of the invention In detail, a protein of the invention is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the protein of the invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be removed and further analyzed. To identify and isolate a binding partner, the entire protein, for instance a protein comprising the entire amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 can be used. Alternatively, a fragment of the protein can be used.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from human kidney tissue, for instance, renal biopsy tissue or tissue culture cells. Alternatively, cellular extracts may be prepared from normal human kidney tissue or available cell lines, particularly kidney derived cell lines.

A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with the protein of the invention under conditions in which association of the protein with the binding partner can

occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins. Alternatively, binding partners may be identified using a Far-Western assay according to the procedures of Takayama *et al.* (1997) Methods Mol. Biol. 69:171-84 or Sauder *et al.* J Gen.Virol. 77(5):991-6 or identified through the use of epitope tagged proteins or GST fusion proteins.

Alternatively, the nucleic acid molecules of the invention can be used in a yeast two-hybrid system. The yeast two-hybrid system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

H. Methods to Identify Agents that Modulate the Expression a Nucleic Acid Encoding the Gene Associated with Renal Compensatory Hypertrophy.

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding a protein of the invention

such as a protein having the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 or SEQ ID NO: 4, if it is capable of up- or down-regulating expression of the nucleic acid in a cell.

In one assay format, cell lines that contain reporter gene fusions between the open reading frame defined by nucleotides 257-1129 of SEQ ID NO: 1, 130-1002 of SEQ ID NO: 3, and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.* (1990) *Anal. Biochem.* 188:245-254). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a protein of the invention such as the protein having SEQ ID NO: 2 or SEQ ID NO: 4. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Laboratory Press, 1989).

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential

probe:non-target hybrids.

Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual,* 2nd Ed. Cold Spring Harbor Laboratory Press, NY, 1989) or Ausubel *et al.* (*Current Protocols in Molecular Biology,* Greene Publishing Co., NY, 1995).

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.* and Ausubel *et al.* as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 or SEQ ID NO: 4 are identified.

Hybridization for qualitative and quantitative analysis of mRNAs may also be carried out by using a RNase Protection Assay (i.e., RPA, see Ma *et al.* (1996) Methods 10: 273-238). Briefly, an expression vehicle comprising cDNA encoding the gene product and a phage specific DNA dependent RNA polymerase promoter (e.g., T7, T3 or SP6 RNA polymerase) is linearized at the 3' end of the cDNA molecule, downstream from the phage promoter, wherein such a linearized molecule is subsequently used as a template for synthesis of a labeled antisense transcript of the cDNA by *in vitro* transcription. The labeled transcript is then hybridized to a mixture of isolated RNA (i.e., total or fractionated mRNA) by incubation at 45°C overnight in a buffer comprising 80% formamide, 40 mM

Pipes, pH 6.4, 0.4 M NaCl and 1 mM EDTA. The resulting hybrids are then digested in a buffer comprising 40 µg/ml ribonuclease A and 2 µg/ml ribonuclease. After deactivation and extraction of extraneous proteins, the samples are loaded onto urea/polyacrylamide gels for analysis.

In another assay format, agents which effect the expression of the instant gene products, cells or cell lines would first be identified which express said gene products physiologically (e.g., see example Figure 1 for tissue distribution via Northern blot, however, RPAs may serve the identical purpose of expression selection). Cell and/or cell lines so identified would be expected to comprise the necessary cellular machinery such that the fidelity of modulation of the transcriptional apparatus is maintained with regard to exogenous contact of agent with appropriate surface transduction mechanisms and/or the cytosolic cascades. Further, such cells or cell lines would be transduced or transfected with an expression vehicle (e.g., a plasmid or viral vector) construct comprising an operable non-translated 5'-promoter containing end of the structural gene encoding the instant gene products fused to one or more antigenic fragments, which are peculiar to the instant gene products, wherein said fragments are under the transcriptional control of said promoter and are expressed as polypeptides whose molecular weight can be distinguished from the naturally occurring polypeptides or may further comprise an immunologically distinct tag. Such a process is well known in the art (see Maniatis).

Cells or cell lines transduced or transfected as outlined above would then be contacted with agents under appropriate conditions; for example, the agent comprises a pharmaceutically acceptable excipient and is contacted with cells comprised in an aqueous physiological buffer such as phosphate buffered saline (PBS) at physiological pH, Eagles balanced salt solution (BSS) at physiological pH, PBS or BSS comprising serum or conditioned media comprising PBS or BSS and/or serum incubated at 37° C. Said conditions may be modulated as deemed necessary by one of skill in the art. Subsequent to contacting the cells with the agent, said cells will be disrupted and the polypeptides of the lysate are fractionated such that a polypeptide fraction is pooled and contacted with an antibody to be further processed by immunological assay (e.g., ELISA, immunoprecipitation or Western blot). The pool of proteins isolated from the "agent-contacted" sample will be compared with a control sample where only the excipient is contacted with the cells and an increase or decrease in the immunologically generated

signal from the "agent-contacted" sample compared to the control will be used to distinguish the effectiveness of the agent.

I. Methods to Identify Agents that Modulate at Least One Activity of the Renal Compensatory Hypertrophy Associated Protein.

Another embodiment of the present invention provides methods for identifying agents that modulate at least one activity of a protein of the invention such as the protein having the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4. Such methods or assays may utilize any means of monitoring or detecting the desired activity.

In one format, the relative amounts of a protein of the invention between a cell population that has been exposed to the agent to be tested compared to an un-exposed control cell population may be assayed. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

Antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins of the invention if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. (Rockford, IL), may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some

applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein (Nature (1975) 256:495-497) or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonals or the polyclonal antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as the Fab, Fab', of $F(ab')_2$ fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin, such as humanized antibodies.

Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a nonrandom basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. As described in the Examples, there are proposed leucine zipper, glycosylation and phosphorylation sites in the protein having SEQ ID NO: 2 or SEQ ID NO: 4. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to or a derivative of a putative leucine zipper.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. Dominant negative proteins, DNAs encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. "Mimic" used herein refers to the modification of a region or several regions of a peptide molecule to provide a structure chemically different from the parent peptide but topographically and functionally similar to the parent peptide (see Grant GA. in: Meyers (ed.) Molecular Biology and Biotechnology (New York, VCH Publishers, 1995), pp. 659-664). A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

Another class of agents of the present invention are antibodies immunoreactive with critical positions of proteins of the invention. Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

J. Uses for Agents that Modulate at Least One Activity of the Proteins.

As provided in the Examples, the proteins and nucleic acids of the invention, such as the protein having the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, are expressed in renal tissue in a compensatory hypertrophy model. Agents that modulate, up-or-down-regulate the expression of the protein or agents such as agonists or antagonists of at least one activity of the protein may be used to modulate biological and pathologic processes associated with the protein's function and activity.

As used herein, a subject can be any mammal, so long as the mammal is in need of modulation of a pathological or biological process mediated by a protein of the invention. The term "mammal" is defined as an individual belonging to the class Mammalia. The invention is particularly useful in the treatment of human subjects.

Pathological processes refer to a category of biological processes which produce a deleterious effect. For example, expression of a protein of the invention may be associated with kidney cell growth regeneration and/or recovery from kidney disease. As described in the Figures and Examples, expression of a protein of the invention may also be associated with tissue remodeling, inflammation, loss of cell cycle control, and the like. As used herein, an agent is said to modulate a pathological process when the agent reduces the degree or severity of the process. For instance, kidney damage or ESRD may be prevented or disease progression modulated by the administration of agents which up-regulate or modulate in some way the expression or at least one activity of a protein of the invention.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular pathological process. For example, an agent of the present invention can be administered in combination with other known drugs or may be combined with dialysis or anti-rejection drugs used during transplantation. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, transdermal, or buccal routes. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

The present invention further provides compositions containing one or more agents which modulate expression or at least one activity of a protein of the invention. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typical dosages comprise 0.1 to 100 μg/kg body wt. The preferred dosages comprise 0.1 to 10 μg/kg body wt. The most preferred dosages comprise 0.1 to 1 μg/kg body wt.

In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active

compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension may also contain stabilizers. Liposomes can also be used to encapsulate the agent for delivery into the cell.

The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral or topical administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

In practicing the methods of this invention, the compounds of this invention may be used alone or in combination, or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be coadministered along with other compounds typically prescribed for these conditions according to generally accepted medical practice. The compounds of this invention can be utilized *in vivo*, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or *in vitro*.

K. Transgenic Animals

Transgenic animals containing mutant, knock-out or modified genes corresponding to the cDNA sequences of SEQ ID NO: 1 or SEQ ID NO: 3 are also included in the invention. Transgenic animals are genetically modified animals into which recombinant, exogenous or cloned genetic material has been experimentally transferred. Such genetic material is often referred to as a "transgene". The nucleic acid sequence of the transgene, in this case a form of SEQ ID NO: 1 or SEQ ID NO: 3, may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene. The transgene may consist of nucleic acid

sequences derived from the genome of the same species or of a different species than the species of the target animal.

The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic animal to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic animals.

The alteration or genetic information may be foreign to the species of animal to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

Transgenic animals can be produced by a variety of different methods including transfection, electroporation, microinjection, gene targeting in embryonic stem cells and recombinant viral and retroviral infection (*see, e.g.*, U.S. Patent No. 4,736,866; U.S. Patent No. 5,602,307; Mullins *et al.* (1993) Hypertension 22(4):630-633; Brenin *et al.* (1997) Surg. Oncol. 6(2)99-110; Tuan (ed.), *Recombinant Gene Expression Protocols*, Methods in Molecular Biology No. 62, Humana Press (1997)).

A number of recombinant or transgenic mice have been produced, including those which express an activated oncogene sequence (U.S. Patent No. 4,736,866); express simian SV40 T-antigen (U.S. Patent No. 5,728,915); lack the expression of interferon regulatory factor 1 (IRF-1) (U.S. Patent No. 5,731,490); exhibit dopaminergic dysfunction (U.S. Patent No. 5,723,719); express at least one human gene which participates in blood pressure control (U.S. Patent No. 5,731,489); display greater similarity to the conditions existing in naturally occurring Alzheimer's disease (U.S. Patent No. 5,720,936); have a reduced capacity to mediate cellular adhesion (U.S. Patent No. 5,602,307); possess a bovine growth hormone gene (Clutter *et al.* (1996) Genetics 143(4):1753-1760); or, are capable of generating a fully human antibody response (McCarthy (1997) The Lancet 349(9049):405).

While mice and rats remain the animals of choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species. Transgenic procedures have been successfully utilized in a variety of non-murine animals, including sheep, goats, pigs, dogs, cats, monkeys, chimpanzees,

hamsters, rabbits, cows and guinea pigs (*see, e.g.*, Kim *et al.* (1997) Mol. Reprod. Dev. 46(4):515-526; Houdebine (1995) Reprod. Nutr. Dev. 35(6):609-617; Petters (1994) Reprod. Fertil. Dev. 6(5):643-645; Schnieke *et al.* (1997) Science 278(5346):2130-2133; and Amoah (1997) J. Animal Science 75(2):578-585).

The method of introduction of nucleic acid fragments into recombination competent mammalian cells can be by any method which favors co-transformation of multiple nucleic acid molecules. Detailed procedures for producing transgenic animals are readily available to one skilled in the art, including the disclosures in U.S. Patent No. 5,489,743 and U.S. Patent No. 5,602,307.

L. Diagnostic Methods

As described in the Examples, the genes and proteins of the invention may be used to diagnose or monitor kidney disease, kidney function, or to track disease progression. For instance, expression of the genes or proteins of the invention can be used to differentiate between Minimal Change disease, IgAN and NCGN. Expression levels can also be used to differentiate between various stages or the severity of IgAN. One means of diagnosing renal disease using the nucleic acid molecules or proteins of the invention involves obtaining kidney tissue from living subjects. Obtaining tissue samples from living sources is problematic for tissues such as kidney. However, due to the nature of the treatment paradigms for kidney patients, biopsy may be necessary. If possible, renal biopsy tissue may be obtained percutaneously.

The use of molecular biological tools has become routine in forensic technology. For example, nucleic acid probes may be used to determine the expression of a nucleic acid molecule comprising all or at least part of the sequences of SEQ ID NO: 1 or SEQ ID NO: 3 in forensic/pathology specimens. Further, nucleic acid assays may be carried out by any means of conducting a transcriptional profiling analysis. In addition to nucleic acid analysis, forensic methods of the invention may target the proteins of the invention, particularly a protein comprising SEQ ID NO: 2 to determine up or down regulation of the genes (Shiverick *et al., Biochim Biophys Acta* (1975) 393(1):124-33).

Methods of the invention may involve treatment of tissues with collagenases or other proteases to make the tissue amenable to cell lysis (Semenov *et al., Biull Eksp Biol Med* (1987) 104(7):113-6). Further, it is possible to obtain biopsy samples from different

regions of the kidney for analysis.

Assays to detect nucleic acid or protein molecules of the invention may be in any available format. Typical assays for nucleic acid molecules include hybridization or PCR based formats. Typical assays for the detection of proteins, polypeptides or peptides of the invention include the use of antibody probes in any available format such as *in situ* binding assays, etc. See Harlow *et al.*, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988. In preferred embodiments, assays are carried-out with appropriate controls.

The above methods may also be used in other diagnostic protocols, including protocols and methods to detect disease states in other tissues or organs, for example the adrenal gland, pancreas, liver, prostate, stomach and testis.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

Example 1

Identification of Differentially Expressed Kidney mRNA

Kidney tissue was obtained from mice in which the other kidney was removed at timepoints over 8 weeks (see Figure 7).

Total cellular RNA was prepared from the kidney tissue described above as well as from control kidney tissue using the procedure of Newburger *et al.*(1981) *J. Biol. Chem.* 266(24): 16171-7 and Newburger *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5215-5219.

Synthesis of cDNA was performed as previously described by Prashar *et al.* in WO 97/05286 and in Prashar *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:659-663.

Briefly, cDNA was synthesized according to the protocol described in the GIBCO/BRL kit for cDNA synthesis. The reaction mixture for first-strand synthesis included 6 μg of total RNA, and 200 ng of a mixture of 1-base anchored oligo(dT) primers with all three possible anchored bases

(ACGTAATACGACTCACTATAGGGCGAATTGGGTCGACTTTTTTTTTTTTTTTT n1 wherein n1=A/C or G) (SEQ ID NO: 6) along with other components for first-strand synthesis reaction except reverse transcriptase. This mixture was incubated at 65°C for 5m, chilled on ice and the process repeated. Alternatively, the reaction mixture may include 10 μg of total RNA, and 2 pmol of 1 of the 2-base anchored oligo(dT) primers with a heel such as RP5.0 (CTCTCAAGGATCTTACCGCTT $_{18}$AT) (SEQ ID NO: 7), or RP6.0 (TAATACCGCGCCACATAGCAT $_{18}$CG) (SEQ ID NO: 8), or RP9.2 (CAGGGTAGACGACGCTACGCT$_{18}$GA) (SEQ ID NO: 9) along with other components for first-strand synthesis reaction except reverse transcriptase. This mixture was then layered with mineral oil and incubated at 65°C for 7 min followed by 50°C for another 7 min. At this stage, 2μl of Superscript® reverse transcriptase (200 units/μl; GIBCO/BRL) was added quickly and mixed, and the reaction continued for 1 hr at 45-50°C. Second-strand synthesis was performed at 16°C for 2 hr. At the end of the reaction, the cDNAs were precipitated with ethanol and the yield of cDNA was calculated. In our experiments, ≈200 ng of cDNA was obtained from 10μg of total RNA.

The adapter oligonucleotide sequences were

A1 (TAGCGTCCGGCGCAGCGACGGCCAG) (SEQ ID NO: 10) and

A2 (GATCCTGGCCGTCGGCTGTCTGTCGGCGC) (SEQ ID NO: 11). One microgram of oligonucleotide A2 was first phosphorylated at the 5′ end using T4 polynucleotide kinase (PNK). After phosphorylation, PNK was heat-denatured, and 1μg of the oligonucleotide A1 was added along with 10× annealing buffer (1 M NaCl/100 mM Tris-HCl, pH8.0/10 mM EDTA, pH8.0) in a final vol of 20 μl. This mixture was then heated at 65°C for 10 min followed by slow cooling to room temperature for 30 min, resulting in formation of the Y adapter at a final concentration of 100 ng/μl. About 20 ng of the cDNA was digested with 4 units of *Bgl* II in a final vol of 10 μl for 30 min at 37°C. Two microliters (≈4 ng of digested cDNA) of this reaction mixture was then used for ligation to 100 ng (≈50-fold) of the Y-shaped adapter in a final vol of 5μl for 16 hr at 15°C. After ligation, the reaction mixture was diluted with water to a final vol of 80 μl

(adapter ligated cDNA concentration, ≈50 pg/μl) and heated at 65°C for 10 min to denature T4 DNA ligase, and 2-μl aliquots (with ≈100 pg of cDNA) were used for PCR.

The following sets of primers were used for PCR amplification of the adapter ligated 3′ -end cDNAs:

TGAAGCCGAGACGTCGGTCG$(T)_{18}$ n1, n2 (wherein n1, n2 = AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG and GT) (SEQ ID NO: 12) as the 3' primer with A1 as the 5' primer or alternatively

RP5.0, RP6.0, or RP9.2 used as 3′ primers with primer A1.1 serving as the 5′ primer. To detect the PCR products on the display gel, 24 pmol of oligonucleotide A1 or A1.1 was 5′-end-labeled using 15 μl of [γ-^{32} P]ATP (Amersham; 3000 Ci/mmol) and PNK in a final volume of 20 μl for 30 min at 37°C. After heat denaturing PNK at 65°C for 20 min, the labeled oligonucleotide was diluted to a final concentration of 2 μM in 80 μl with unlabeled oligonucleotide A1.1. The PCR mixture (20μl) consisted of 2 μl (≈100 pg) of the template, 2μl of 10× PCR buffer (100 mM Tris·HCl, pH 8.3/500 mM KCl), 2 μl of 15 mM $MgCl_2$ to yield 1.5 mM final Mg^{2+} concentration optimum in the reaction mixture, 200 μM dNTPs, 200 nM each 5′ and 3′ PCR primers, and 1 unit of Amplitaq Gold®. Primers and dNTPs were added after preheating the reaction mixture containing the rest of the components at 85°C. This "hot start" PCR was done to avoid artefactual amplification arising out of arbitrary annealing of PCR primers at lower temperature during transition from room temperature to 94°C in the first PCR cycle. PCR consisted of 5 cycles of 94°C for 30 sec, 55°C for 2 min, and 72°C for 60 sec followed by 25 cycles of 94°C for 30 sec, 60°C for 2 min, and 72°C for 60 sec. A higher number of cycles resulted in smeary gel patterns. PCR products (2.5μl) were analyzed on 6% polyacrylamide sequencing gel. For double or multiple digestion following adapter ligation, 13.2 μl of the ligated cDNA sample was digested with a secondary restriction enzyme(s) in a final vol of 20 μl. From this solution, 3μl was used as template for PCR. This template vol of 3 μl carried ≈ 100 pg of the cDNA and 10 mM $MgCl_2$ (from the 10× enzyme buffer), which diluted to the optimum of 1.5 mM in the final PCR vol of 20 μl. Since Mg^{2+} comes from the restriction enzyme buffer, it was not included in the reaction mixture when amplifying secondarily cut cDNA. Individual cDNA fragments corresponding to mRNA species were separated by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography. Bands were extracted from the display gels as described by Liang *et al.* (1995 *Curr. Opin.*

Immunol. 7:274-280), reamplified using the 5′ and 3′ primers, and subcloned into pCR-Script with high efficiency using the PCR-Script® cloning kit from Stratagene. Plasmids were sequenced by cycle sequencing on an automated sequencer. Alternatively, bands were extracted (cored) from the display gels, PCR amplified and sequenced directly without subcloning.

A 391 bp fragment was identified in tissue samples from the compensatory hypertrophy model which corresponds to a mRNA species which was up-regulated compared to control kidney tissue starting at 4 days post excision of one kidney. Expression of this mRNA species was up-regulated compared to control kidney tissue at four weeks post-excision. The fragment corresponding to the band on the gel was sequenced (see Figure 7). The sequence of the band is:

ggtacaaggc ctgcatcaag taccctgagt ggaagcagaa acaccagcca catttcaagc 60
catggctgca cccggagcag agcccgttac ccagcctggc gctgtcagag ctgtcggtgc 120
aacacgcaga ctcactggag aacatcgacg agagcgccgt gacagagagc cgggaggagc 180
gggcgggtgg agcaggagag ggcagcgatg acgacacctc gctcacctga gaaccgcctc 240
tcttcaagga ccaagacagg actggggtgg aaccctgggg cactgagcct tctgcacttc 300
ctcccttccc ccacctgcct tttgggggca ctgggctcct gctcaggtgg ctgcggcatg 360
gctggacatg gccccaataa atgaaccaca c 391 (SEQ ID NO: 5)

Example 2

Cloning of a Full Length Human and Mouse cDNA Corresponding to the 391 bp Fragment

The full length human cDNA (SEQ ID NO: 1) and murine cDNA (SEQ ID NO: 3) corresponding to the 391 bp fragment obtained in Example 1 (SEQ ID NO: 5) were obtained by the oligo-pulling method. Briefly, a gene-specific oligo was designed based on the sequence of SEQ ID NO: 5. The oligo was labeled with biotin and used to hybridize with 2 μg of single strand plasmid DNA (cDNA recombinants) from a human kidney cDNA library following the procedures of Sambrook *et al.* The hybridized cDNAs were separated by streptavidin-conjugated beads and eluted by heating. The eluted cDNA was converted to double strand plasmid DNA and used to transform *E. coli* cells (DH10B)

and the longest cDNA was screened. After confirmation by PCR using gene-specific primers, the cDNA clone was subjected to DNA sequencing.

The nucleotide sequence of the full-length human cDNA corresponding to the differentially regulated band having the sequence of SEQ ID NO: 5 is set forth in SEQ ID NO: 1. The cDNA comprises 1164 base pairs with an open reading frame encoding a protein of 291 amino acids. The amino acid sequence is presented in SEQ ID NOs: 1 and 2.

A gene-specific oligo was also designed based on the sequence of SEQ ID NO: 5 to screen a mouse cDNA library. The nucleotide sequence of the full-length murine cDNA corresponding to the differentially regulated band having the sequence of SEQ ID NO: 5 is set forth in SEQ ID NO: 3. The cDNA comprises 1173 base pairs with an open reading frame encoding a protein of 291 amino acids. The amino acid sequence is presented in SEQ ID nos: 3 and 4.

Bestfit analysis of the amino acid alignment between the human (SEQ ID NO: 2) and murine (SEQ ID NO: 4) amino acid sequences indicates 97% identity (see Figure 1). The predicted hydrophilicity of peptide fragments within SEQ ID NO: 2 and SEQ ID NO: 4 are shown in Figure 2A and B, respectively. Analysis of the amino acid sequence of SEQ ID NO: 2 and SEQ ID NO: 4 predicts the following regions homologous with known protein motifs:

1) regions resembling a leucine zipper are present at amino acids 220-252;

2) a predicted N-glycosylation site is at amino acids 44-47 and 147-150;

3) predicted casein kinase II phosphorylation sites are at amino acids 35-38 and 81-84, 271-275 and 284-289;

4) predicted N-myristoylation sites are found at amino acids 277-282 and 280-285, as well as 12-17 in the murine protein; and

5) predicted protein kinase C phosphorylation sites at amino acids 32-34, 118-120, 132-134 and 149-151.

Example 3

Northern blot and PCR Expression Analysis

The tissue distribution of RNA encoding the protein of SEQ ID NO: 2 was

analyzed by Northern blot (see Figure 3) as well as PCR expression analysis of RNA isolated from various tissues (see Figure 4). RNA was isolated from human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas using standard procedures. Northern blots were prepared using a probe derived from SEQ ID NO: 1 with hybridization conditions as described by Sambrook *et al.* (Church-Gilbert Hybridization buffer ; hybridized at 65°C overnight; wash at 65°C until 0.5 x SSC with 0.1% SDS; Expose 15 hours at -80°C.). PCR expression analysis was also performed using primers derived from SEQ ID NO: 1 using AmpliTaq Gold PCR® amplification kits (Perkin Elmer).

Quantitative PCR Analysis of Expression Levels

Figures 5 to 6 and 8 to 10 show the results of the quantitative PCR analysis of expression levels of mRNA corresponding to SEQ ID NO: 1 in various human tissue samples. Real time PCR detection was accomplished by the use of the ABI PRISM 7700 Sequence Detection System. The 7700 measures the fluorescence intensity of the sample each cycle and is able to detect the presence of specific amplicons within the PCR reaction. Each sample was assayed for the level of GAPDH and mRNA corresponding to SEQ ID NO: 1. GAPDH detection was performed using Perkin Elmer part#402869 according to the manufacturer's directions. Primers were designed from SEQ ID NO: 1 using Primer Express, a program developed by PE to efficiently find primers and probes for specific sequences. These primers were used in conjunction with SYBR green (Molecular Probes), a nonspecific double stranded DNA dye, to measure the expression level mRNA corresponding to SEQ ID NO: 1, which was normalized to the GAPDH level in each sample.

Differentially Expressed mRNA in Renal Disease

The genes and proteins of the invention can be used to distinguish among various types of renal disease diagnostically and prognostically, or to track disease progress. For example (see Figure 5), a band corresponding to SEQ ID NO: 1 is upregulated in Minimal Change and IgA Nephropathy (IgAN), but down regulated in NCGN; in IgAN it is upregulated in mild and moderate cases, but is expressed at levels comparable to non-

diseased, normal individuals in severe IgAN.

Differentially Expressed mRNA in Renal and non-Renal Tissues

Tissue distribution analysis indicates that the mRNA of the sequence of SEQ ID NO: 1 is widely distributed. For example (see Figure 6), high levels of expression are observed in pancreas, liver, stomach, testes, adrenal glands, prostate, and other tissues, including kidney. This pattern of widespread tissue indicates that the protein of the invention may play a role in any disease anywhere in the body in which inflammation and tissue remodeling occurs, and that the proteins, antibodies an nucleic acids of the invention may be used to diagnosis such indications.

Differentially Expressed mRNA at various Stages of the Cell Cycle

Functionally, expression of the gene of the invention is up-regulated in compensatory hypertrophy following unilateral nephrectomy in the rat (see Figure 7). Up-regulation is also observed in cultured rat mesangial cells under conditions of growth arrest due to serum starvation (Figure 8), while downregulation follows the re-addition of serum to growth factor depleted cells (Figure 9). Upregulation is also observed when human umbilical vein and arterial cells are treated with tumor necrosis factor, TNF-α (Figure 10). These results suggest that suppression of the protein of the invention is necessary for loss of growth control.

Overexpression of the protein of the invention in mesangial cells results in increased cell death (see Figure 11) and a shift from S phase into G2 phase (see Figure 12), showing the biological effects of exogenously modulating its expression.

For Figures 11-13, the expression vector for the protein of SEQ ID NO: 2 was constructed with a CMV-promoter to facilitate transfection and mRNA expression, and designated vector "GN-01". Cells were transfected with the human GN-01 vector or GN-01* mutant vector and protein was detected with anti-FLAG-FITC-labeled antibodies (Figures 11 and 12) . Cell cycle and cell death analysis was performed using flow cytometry to measure BrdU uptake as detected by anti-BrdU-FITC antibodies and by standard gating methods (Figures 8, 9, and 11) or to measure DNA content with

Propridium (Figure 12).

Example 4

Cellular Localization of the Protein

The localization of the protein of SEQ ID NO: 1 within a cell appears to vary by cell type and/or other factors of cell cycle/activation. For example (see Figure 13), in mesangial cells, the expressed protein can be nuclear or excluded from the nucleus. Confocal microscopy was performed using standard techniques.

The above-described properties of the gene and protein of the invention indicate that therapeutics designed to treat renal or other diseases with inflammatory, cell proliferation, and/or tissue remodeling effects may target specific motifs, or the level of expression of the gene or its protein. The overexpression effects leading to increased cell death (Figure 11) and the differential localization of the protein in distinct cell conditions (Figure 13) indicate properties that could be used for designing specific therapeutic treatments. For example, the overexpression or therapeutic regulation of the RNA message of SEQ ID NO: 1 or the protein of SEQ ID NO: 2 could be used to selectively target subpopulations of cells in renal or other diseases resulting in destruction, the prevention of destruction, enhanced growth or mitosis, or other desired modulations of cellular function.

Although the present invention has been described in detail with reference to examples above, it is understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims. All cited patents and publications referred to in this application are herein incorporated by reference in their entirety.

WHAT IS CLAIMED:

1. An isolated nucleic acid molecule selected from the group consisting of: (a) an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4; (b) an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NO: 2 or SEQ ID NO: 4; (c) an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO: 1; (d) an isolated nucleic acid molecule which hybridizes to a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4; and (e) an isolated nucleic acid molecule that encodes a protein that is expressed in compensatory hypertrophy and that exhibits at least about 60% nucleotide sequence identity over nucleotides 257-1129 of SEQ ID NO: 1 or nucleotides 130-1002 of SEQ ID NO: 3 .

2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 257-1129 of SEQ ID NO: 1.

3. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule consists of nucleotides 257-1129 of SEQ ID NO: 1.

4. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 130-1002 of SEQ ID NO: 3.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule consists of nucleotides 130-1002 of SEQ ID NO: 3.

6. The isolated nucleic acid molecule of any one of claims 1-5, wherein said nucleic acid molecule is operably linked to one or more expression control elements.

7. A vector comprising an isolated nucleic acid molecule of any one of claims 1-5.

8. A host cell transformed to contain the nucleic acid molecule of any one of

claims 1-5.

9. A host cell comprising a vector of claim 7.

10. A host cell of claim 9, wherein said host is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

11. A method for producing a polypeptide comprising the step of culturing a host cell transformed with the nucleic acid molecule of any one of claims 1-5 under conditions in which the protein encoded by said nucleic acid molecule is expressed.

12. The method of claim 11, wherein said host cell is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

13. An isolated polypeptide produced by the method of claim 11.

14. An isolated polypeptide or protein selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, an isolated polypeptide comprising a fragment of at least 6 amino acids of SEQ ID NO: 2 or SEQ ID NO: 4, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2 or SEQ ID NO: 4, naturally occurring amino acid sequence variants of SEQ ID NO: 2 or SEQ ID NO: 4 and an isolated polypeptide exhibiting at least about 50% amino acid sequence identity with SEQ ID NO: 2.

15. An isolated antibody that binds to a polypeptide of either claim 13 or 14.

16. The antibody of claim 15 wherein said antibody is a monoclonal or polyclonal antibody.

17. A method of identifying an agent which modulates the expression of a nucleic acid encoding a protein of claim 14, comprising the steps of:

exposing cells which express the nucleic acid to the agent; and

determining whether the agent modulates expression of said nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

18. A method of identifying an agent which modulates at least one activity of a protein of claim 14, comprising the steps of:

exposing cells which express the protein to the agent;

determining whether the agent modulates at least one activity of said protein, thereby identifying an agent which modulates at least one activity of the protein.

19. The method of claim 18, wherein the agent modulates one activity of the protein.

20. A method of identifying binding partners for a protein of claim 14, comprising the steps of:

exposing said protein to a potential binding partner; and

determining if the potential binding partner binds to said protein, thereby identifying binding partners for the protein.

21. A method of modulating the expression of a nucleic acid encoding a protein of claim 14, comprising the step of:

administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein.

22. A method of modulating at least one activity of a protein of claim 14, comprising the step of:

administering an effective amount of an agent which modulates at least one activity of the protein.

23. A transgenic non-human animal modified to contain a nucleic acid molecule of any of claims 1-5.

24. The transgenic animal of claim 23, wherein the nucleic acid molecule contains a mutation that prevents expression of the encoded protein.

25. A method of diagnosing a disease state in a subject, comprising the step of determining the level of expression of a nucleic acid molecule or protein of any one of claims 1-5 or 14.

26. The method of claim 25, wherein the disease state is a renal disease, inflammatory disease or disease associated with cell proliferation.

FIG. 1

```
HUMAN   1   MEKLAASTEPQGPRPVLGRESVQVPDDQDFRSFRSECEAEVGWNLTYSRAGVSVWVQAVE 60
            MEK AASTEPQG RP LGRESVQVPDDQDFRSFRSECEAEVGWNLTYS+AGVSVWVQAVE
MOUSE   1   MEKPAASTEPQGSRPALGRESVQVPDDQDFRSFRSECEAEVGWNLTYSKAGVSVWVQAVE 60

        61  MDRTLHKIKCRMECCDVPAETLYDVLHDIEYRKKWDSNVIETFDIARLTVNADVGYYSWR 120
            MDRTLHKIKCRMECCDVPAETLYDVLHDIEYRKKWDSNVIETFDIARLTVNADVGYYSWR
        61  MDRTLHKIKCRMECCDVPAETLYDVLHDIEYRKKWDSNVIETFDIARLTVNADVGYYSWR 120

        121 CPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPPRKDLVRAVSIQTGYLIQSTGPKS 180
            CPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPPRKDLVRAVSIQTGYLIQSTGPKS
        121 CPKPLKNRDVITLRSWLPMGADYIIMNYSVKHPKYPPRKDLVRAVSIQTGYLIQSTGPKS 180

        181 CVITYLAQVDPKGSLPKWVVNKSSQFLAPKAMKKMYKACLKYPEWKQKHLPHFKPWLHPE 240
            CVITYLAQVDPKGSLPKWVVNKSSQFLAPKAMKKMYKAC+KYPEWKQKH PHFKPWLHPE
        181 CVITYLAQVDPKGSLPKWVVNKSSQFLAPKAMKKMYKACIKYPEWKQKHQPHFKPWLHPE 240

        241 QSPLPSLALSELSVQHADSLENIDESAVAESREERMGGAGGEGSDDDTSLT 291
            QSPLPSLALSELSVQHADSLENIDESAV ESREER GGAGGEGSDDDTSLT
        241 QSPLPSLALSELSVQHADSLENIDESAVTESREERAGGAGGEGSDDDTSLT 291
```

FIG. 2

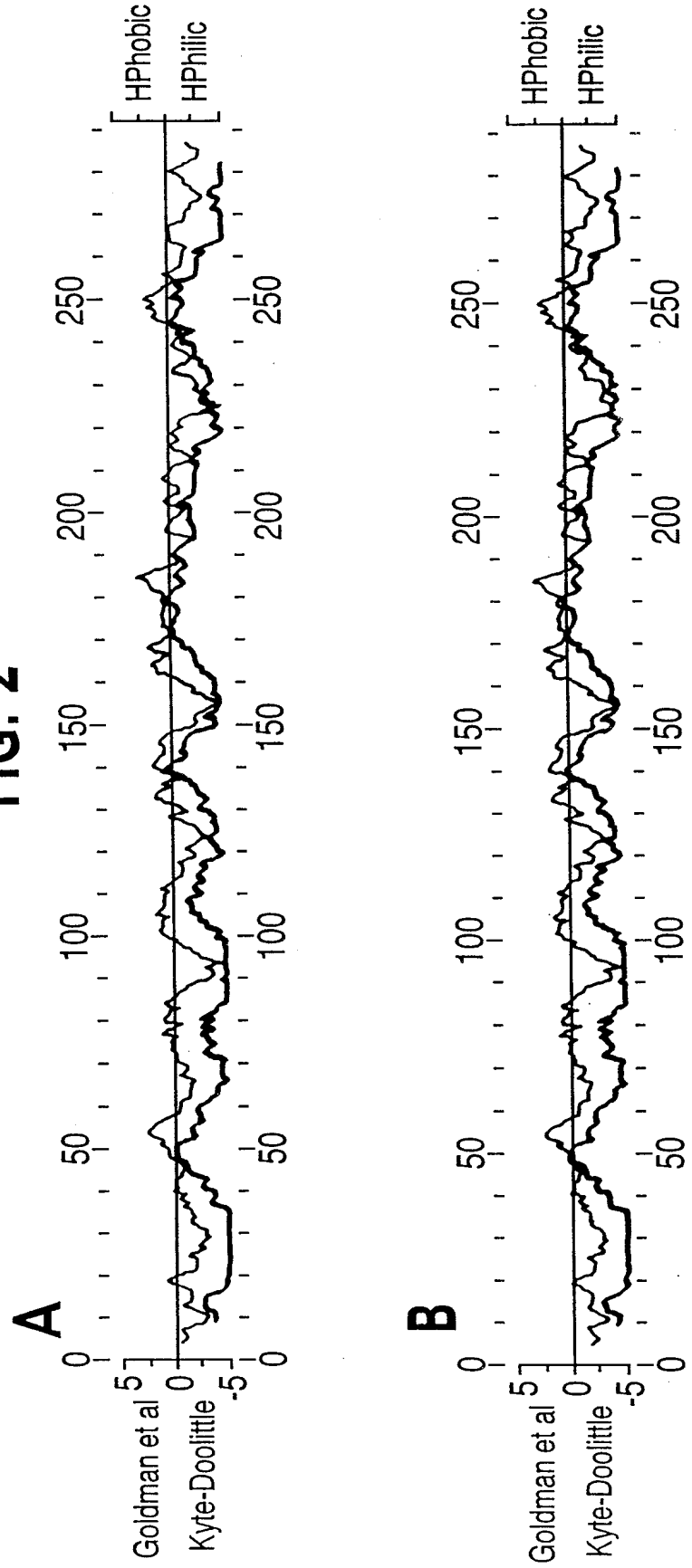

A
B
HPhobic
HPhilic
0
50
100
150
200
250
5
0
-5
Goldman et al
Kyte-Doolittle

FIG. 3

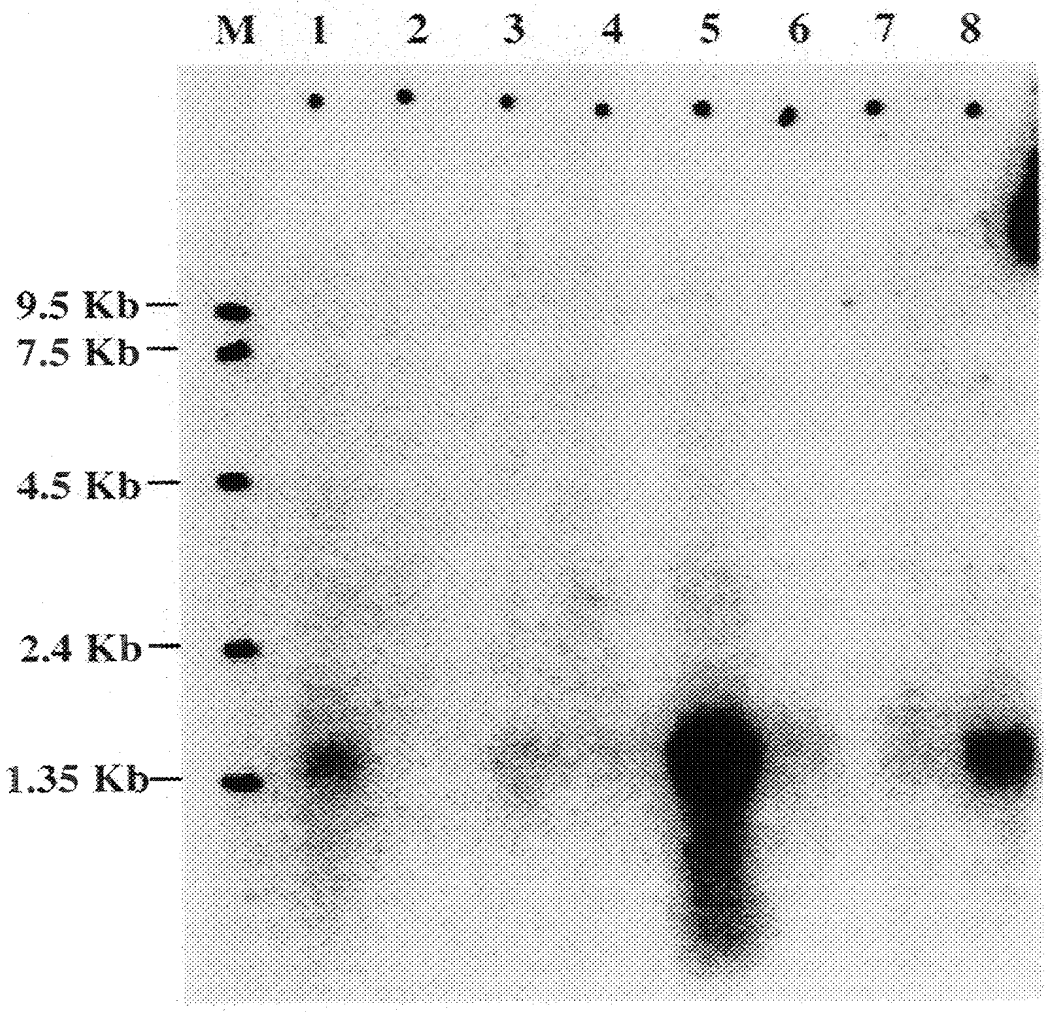


FIG. 4

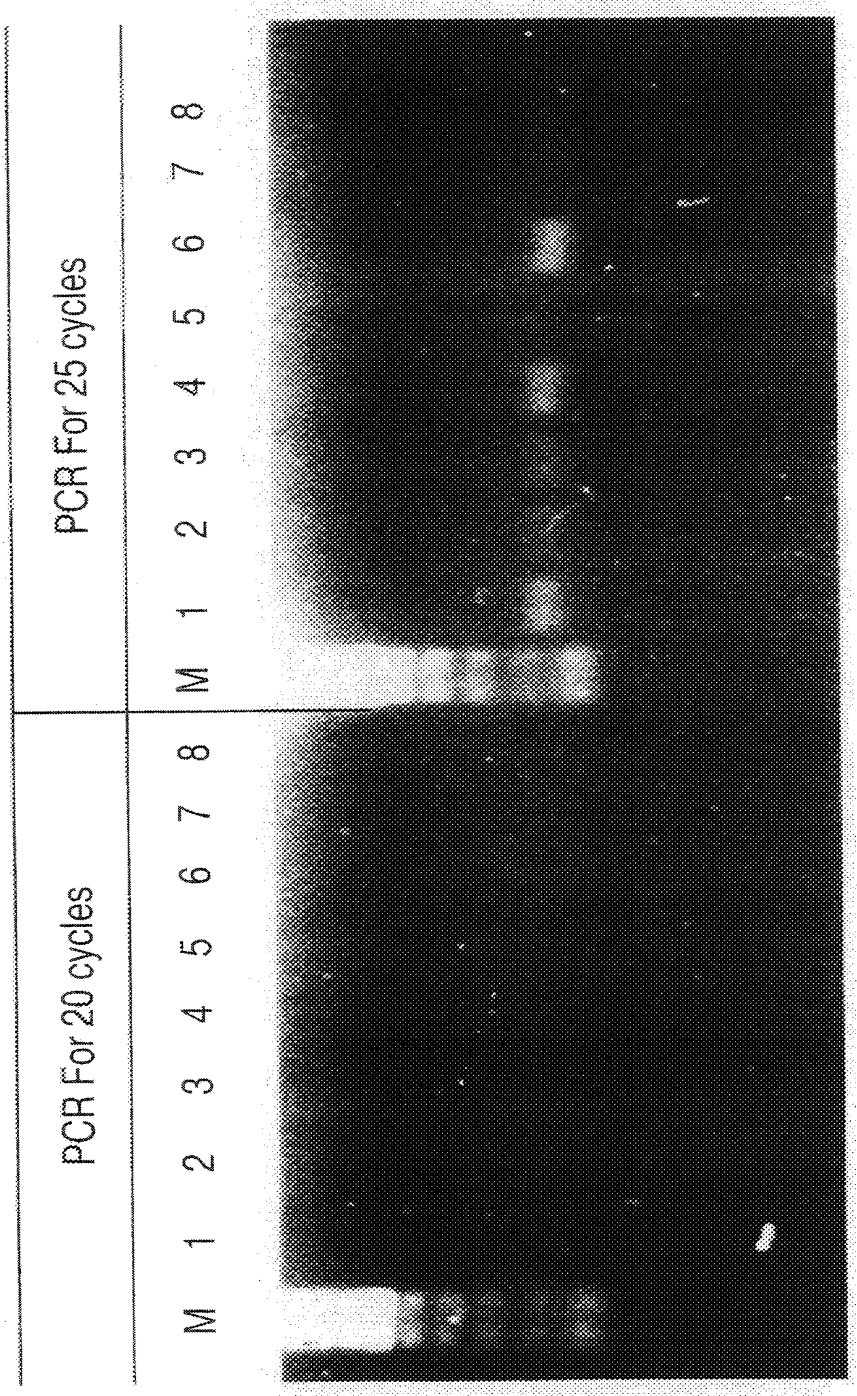


FIG. 5

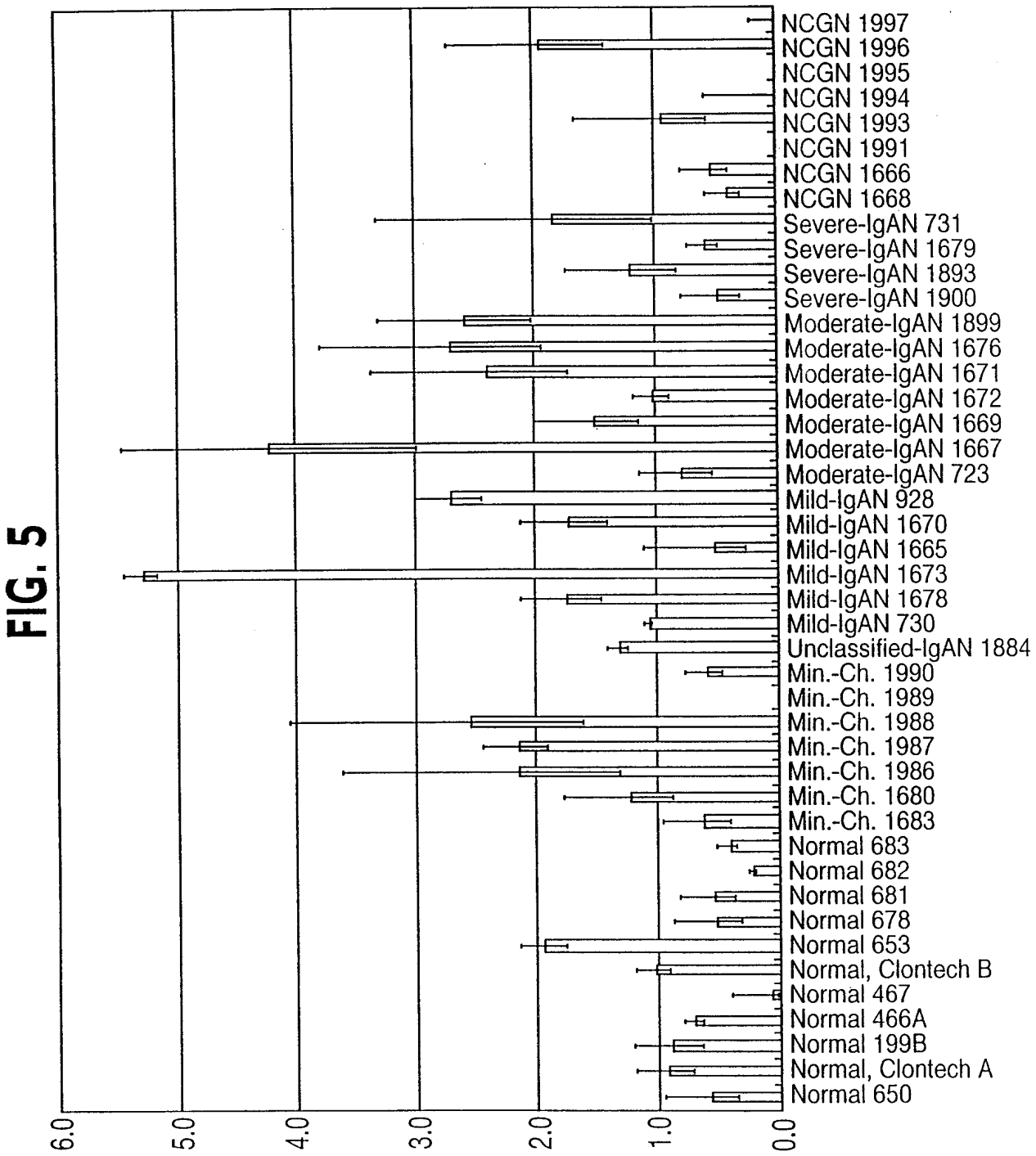

NCGN 1997
NCGN 1996
NCGN 1995
NCGN 1994
NCGN 1993
NCGN 1991
NCGN 1666
NCGN 1668
Severe-IgAN 731
Severe-IgAN 1679
Severe-IgAN 1893
Severe-IgAN 1900
Moderate-IgAN 1899
Moderate-IgAN 1676
Moderate-IgAN 1671
Moderate-IgAN 1672
Moderate-IgAN 1669
Moderate-IgAN 1667
Moderate-IgAN 723
Mild-IgAN 928
Mild-IgAN 1670
Mild-IgAN 1665
Mild-IgAN 1673
Mild-IgAN 1678
Mild-IgAN 730
Unclassified-IgAN 1884
Min.-Ch. 1990
Min.-Ch. 1989
Min.-Ch. 1988
Min.-Ch. 1987
Min.-Ch. 1986
Min.-Ch. 1680
Min.-Ch. 1683
Normal 683
Normal 682
Normal 681
Normal 678
Normal 653
Normal, Clontech B
Normal 467
Normal 466A
Normal 199B
Normal, Clontech A
Normal 650
6.0
5.0
4.0
3.0
2.0
1.0
0.0

FIG. 6

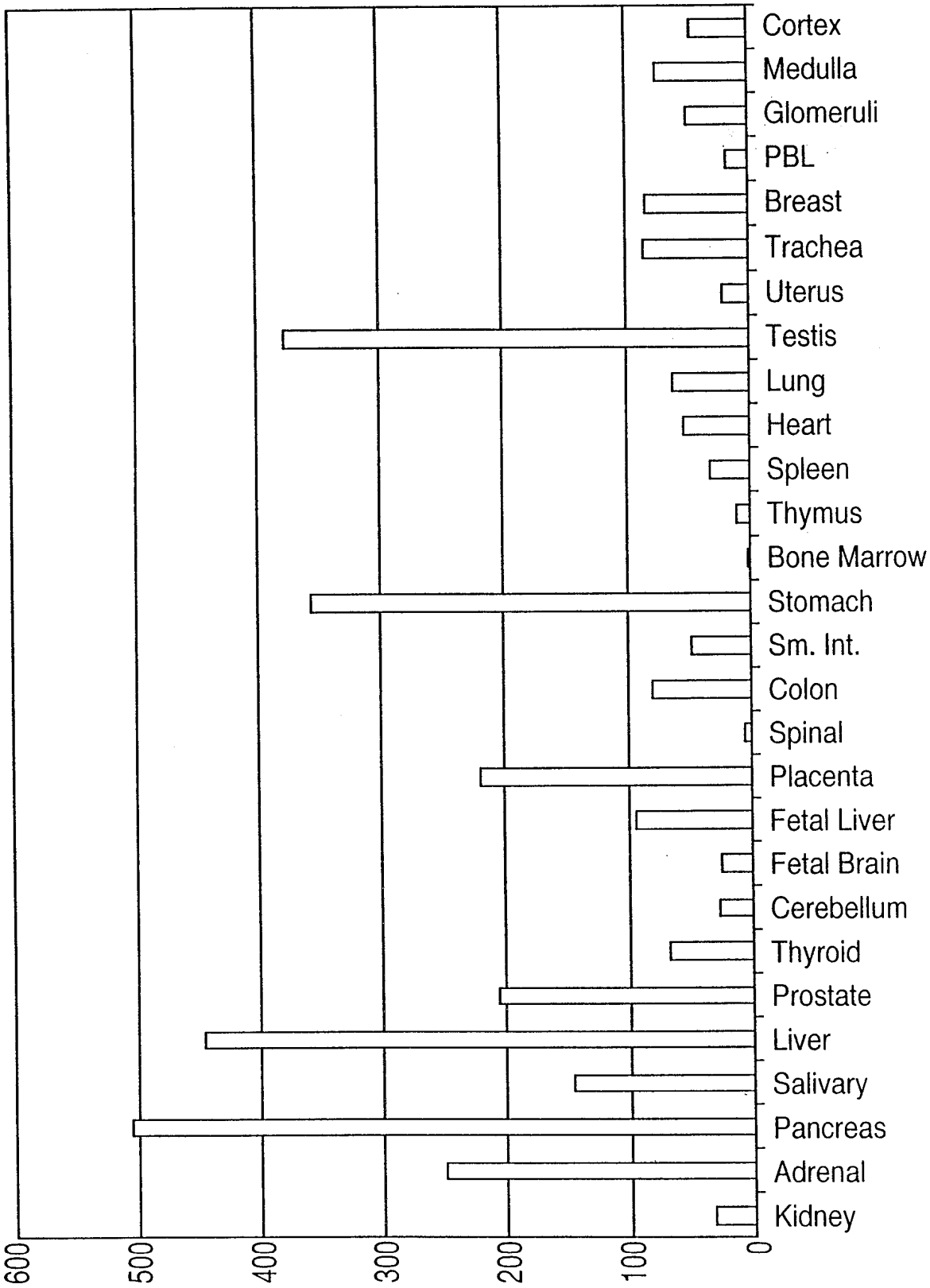

Cortex
Medulla
Glomeruli
PBL
Breast
Trachea
Uterus
Testis
Lung
Heart
Spleen
Thymus
Bone Marrow
Stomach
Sm. Int.
Colon
Spinal
Placenta
Fetal Liver
Fetal Brain
Cerebellum
Thyroid
Prostate
Liver
Salivary
Pancreas
Adrenal
Kidney
600
500
400
300
200
100
0

FIG. 7

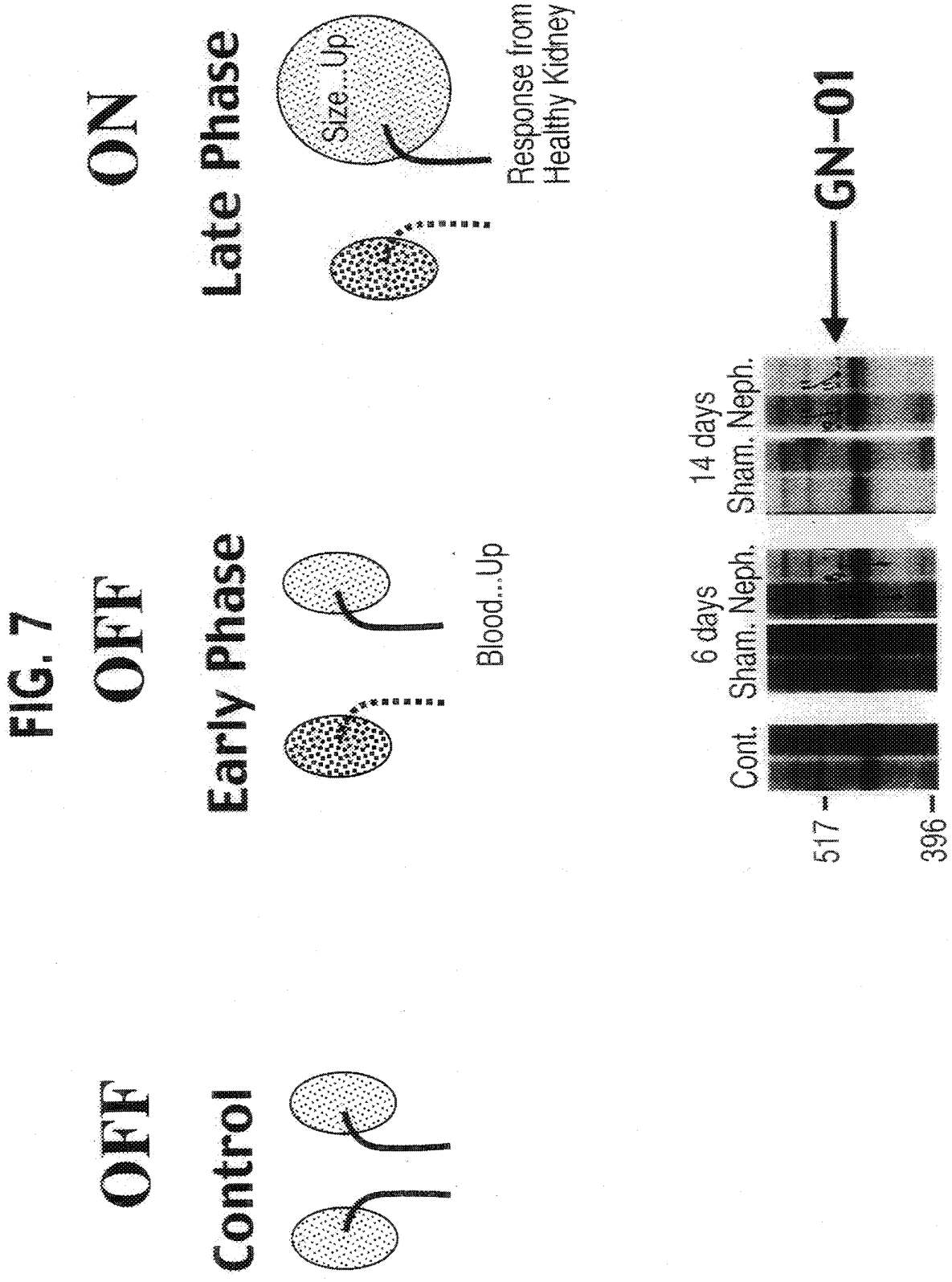

OFF
Control
OFF
Early Phase
Blood...Up
ON
Late Phase
Size...Up
Response from Healthy Kidney
GN-01
6 days
Cont.
Sham.
Neph.
14 days
Sham.
Neph.
517
396

FIG. 8

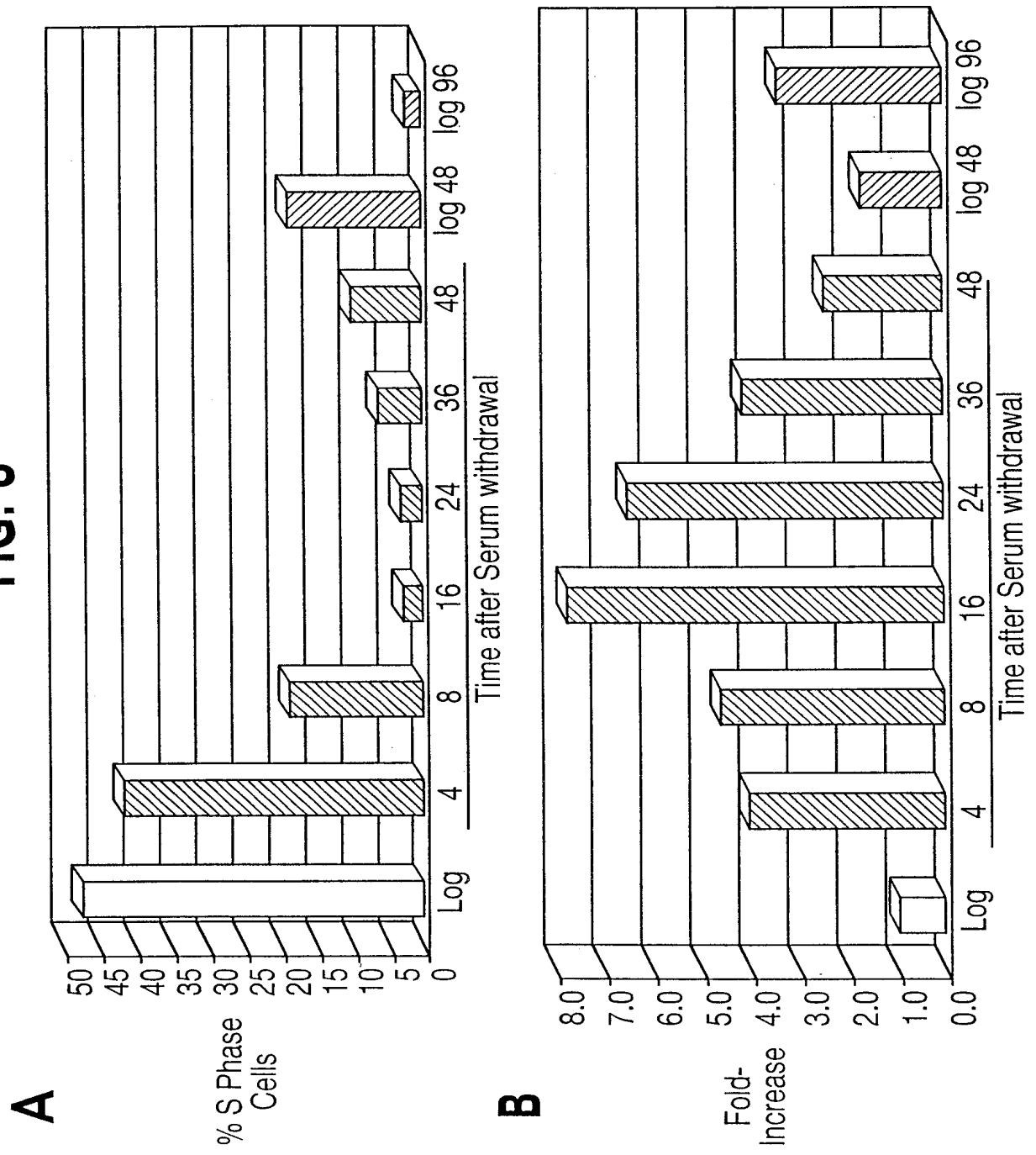

A
% S Phase Cells
50
45
40
35
30
25
20
15
10
5
0
Log
4
8
16
24
36
48
log 48
log 96
Time after Serum withdrawal
B
Fold-Increase
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Log
4
8
16
24
36
48
log 48
log 96
Time after Serum withdrawal

FIG. 9

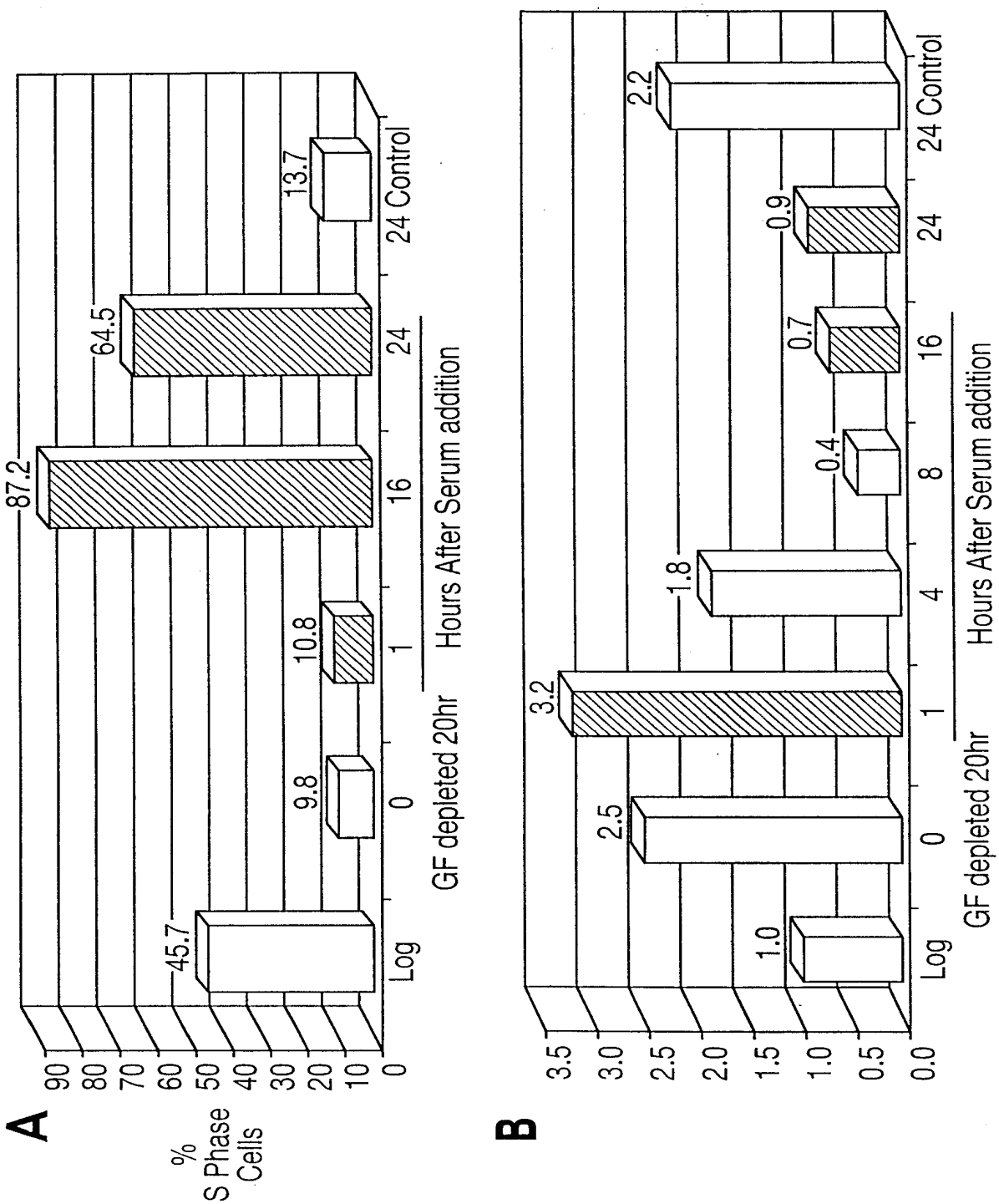

A
% S Phase Cells
0
10
20
30
40
50
60
70
80
90
45.7
9.8
10.8
87.2
64.5
13.7
Log
0
1
16
24
24 Control
GF depleted 20hr
Hours After Serum addition
B
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
1.0
2.5
3.2
1.8
0.4
0.7
0.9
2.2
Log
0
1
4
8
16
24
24 Control
GF depleted 20hr
Hours After Serum addition

FIG. 10

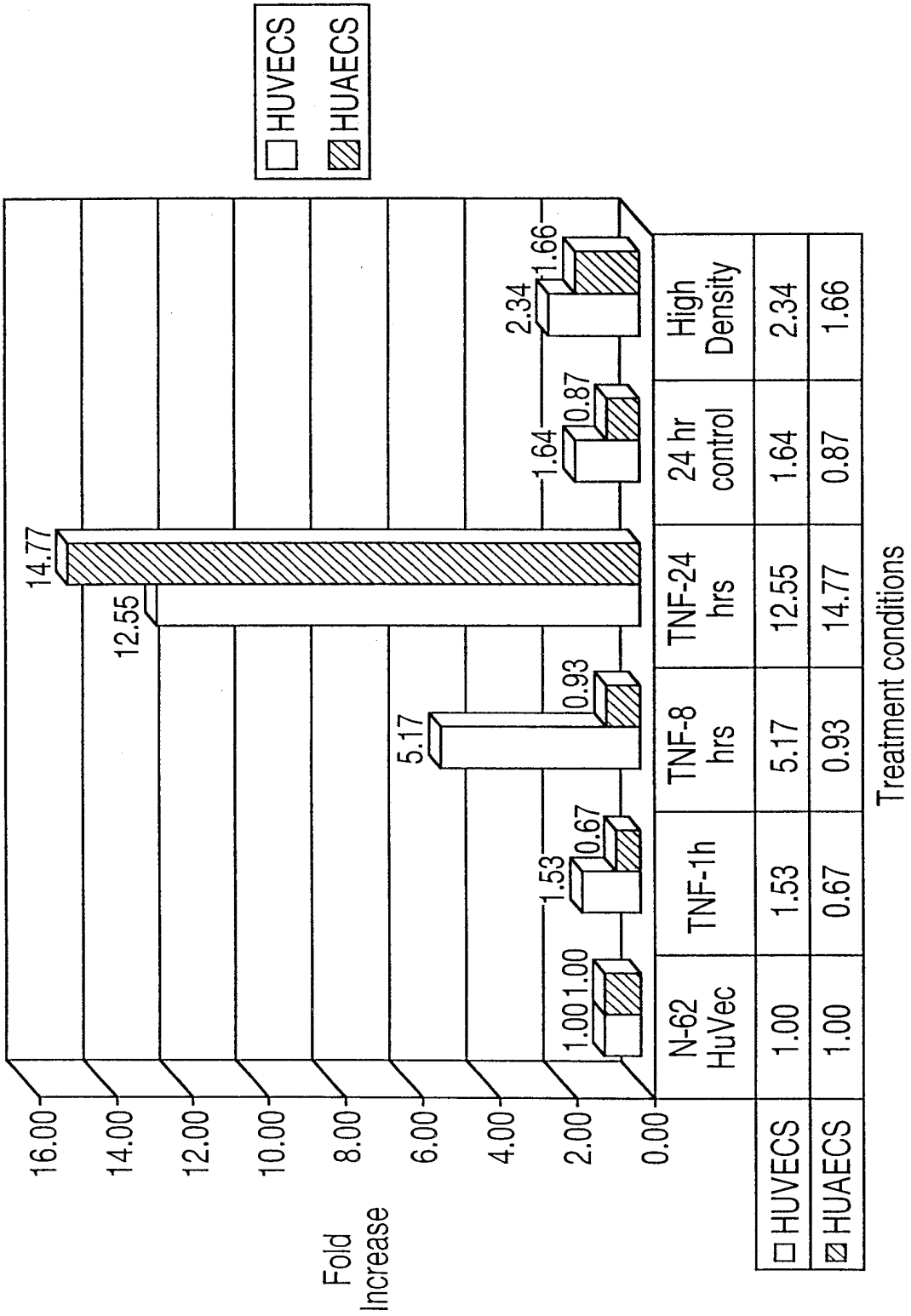

HUVECS
HUAECS
Fold Increase
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
14.77
12.55
5.17
0.93
1.53
0.67
1.00 1.00
2.34
1.66
1.64
0.87
Treatment conditions
N-62 HuVec
TNF-1h
TNF-8 hrs
TNF-24 hrs
24 hr control
High Density
HUVECS 1.00 1.53 5.17 12.55 1.64 2.34
HUAECS 1.00 0.67 0.93 14.77 0.87 1.66

FIG. 11

A

Vector	%G1	%S	%G2	% Death
LacZ	55	28	16	23
GN-01 *	61	23	16	31
GN-01	61	24	15	44

B

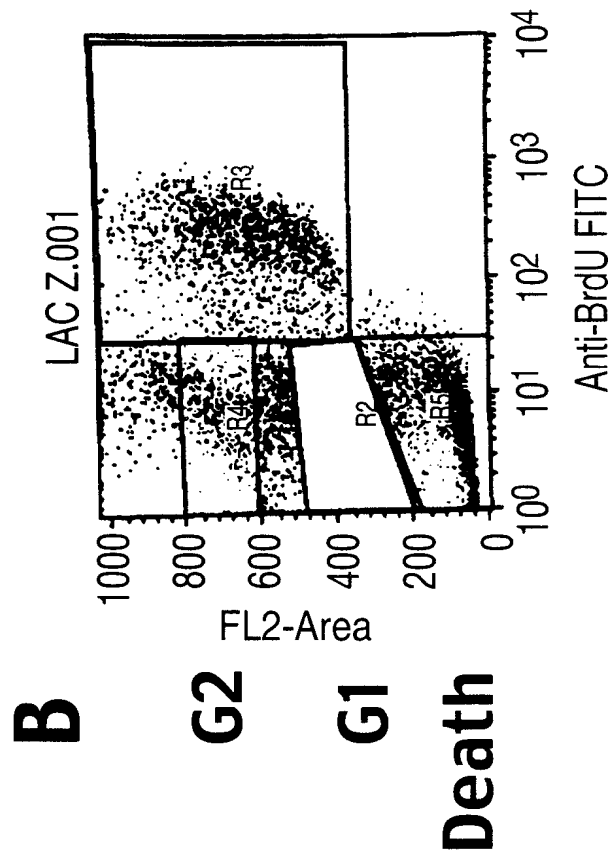
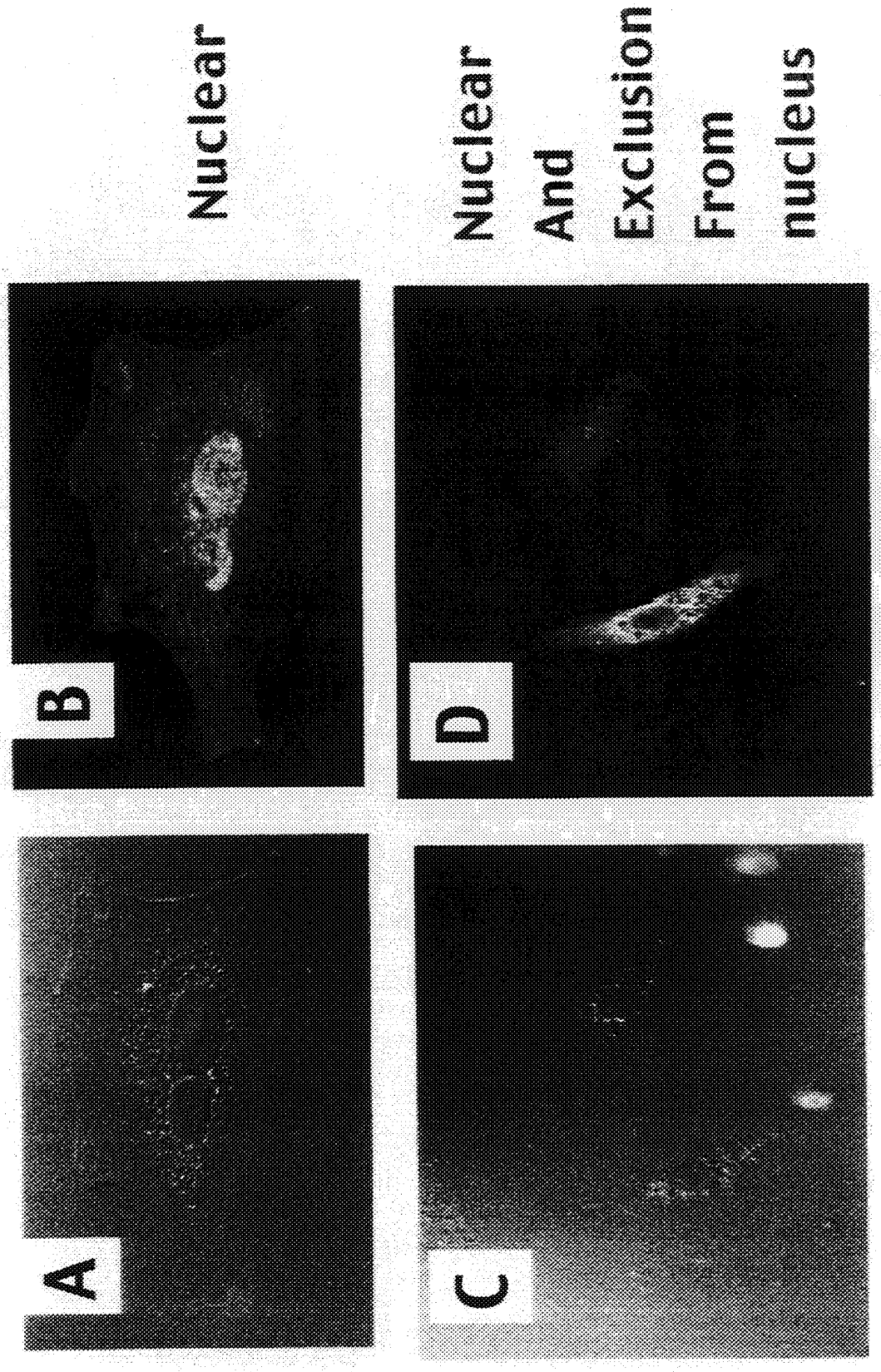


FIG. 12

		Cell Cycle Distribution		
		G1(%)	S (%)	G2(%)
LacZ	Control	43	32	25
GN01	GN01 +	50	18	32
	Control	49	33	18
GN01*	GN01 +	52	27	21
-mutant	Control	52	29	19

FIG. 13


A
B
Nuclear
C
Contrast Image
D
Nuclear And Exclusion From nucleus
GN-01

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05712

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 435/7.1; 435/69.1; 435/325+; 800/8+; 530/350+; 424 130.1+

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAPLUS, MEDLINE, BIOSIS, EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TOMASHEFSKY ET AL. Macromolecular metabolism in renal compensatory hypertrophy. II. Protein turnover. Laboratory Investigation. 1970, Vol. 23, No. 2, pages 190-195, see entire document.	1-26
Y	STEPHAN ET AL. Impairment of renal compensatory hypertrophy by hypothyroidism in the rat. Life Sciences. 1982, Vol. 30, Nos. 7 & 8, pages 623-31, see entire document.	1-26
Y	LING ET AL. Suppressing role of transforming growth factor-β1 on cathepsin activity in cultured kidney tubule cells. Am. J. Physiol. 1995, Vol. 269, pages F911-F917, see entire document.	1-26

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:
"A"	document defining the general state of the art which is not considered to be of particular relevance
"E"	earlier document published on or after the international filing date
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O"	document referring to an oral disclosure, use, exhibition or other means
"P"	document published prior to the international filing date but later than the priority date claimed
"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&"	document member of the same patent family

Date of the actual completion of the international search

24 MAY 2000

Date of mailing of the international search report

20 JUN 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

GAI (JENNIFER) MI LEE

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05712

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SCHWARTZ ET AL. Reversible compensatory hypertrophy in rat kidneys: morphometric characterization. Kidney International 1993, Vol. 43, No. 3, pages 610-614, see entire document.	1-26
A	HILL, J. M. Methylglyoxal bid (guanylhydrazone) inhibits the accumulation of kidney proteins during compensatory renal hypertrophy. Biochem. Biophys. Res. Commun. 1975, Vol. 65, No. 3, pages 961-967, see entire document.	1-26

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05712

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C07H 21/02; C07K 1/00; G01N 33/53; C12P 21/06; C12N 5/00; A61K 39/395; A01K 67/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.1; 435/7.1; 435/69.1; 435/325+; 800/8+; 530/350+; 424 130.1+